Proximity sensing with wavelet generated video

Steven E. Noel
The Center for Advanced Computer Studies
University of Southwestern Louisiana
Lafayette, Louisiana 70504
Email: snoel@cacs.usl.edu

Harold H. Szu
Naval Surface Warfare Center Dahlgren Division
Systems Research and Technology Department
17320 Dahlgren Road
Dahlgren, VA 22448-5100
Email: hszu@nswc.navy.mil

Abstract. *In this paper we introduce wavelet video processing of proximity sensor signals. Proximity sensing is required for a wide range of military and commercial applications, including weapon fuzing, robotics, and automotive collision avoidance. While our proposed method temporarily increases signal dimension, it eventually performs data compression through the extraction of salient signal features. This data compression in turn reduces the necessary complexity of the remaining computational processing. We demonstrate our method of wavelet video processing via the proximity sensing of nearby objects through their Doppler shift. In doing this we perform a continuous wavelet transform on the Doppler signal, after subjecting it to a time varying window. We then extract signal features from the resulting wavelet video, which we use as input to pattern recognition neural networks. The networks are trained to estimate the time varying Doppler shift from the extracted features. We test the estimation performance of the networks, using different degrees of nonlinearity in the frequency shift over time and different levels of noise. We give the analytical result that the signal-to-noise enhancement of our proposed method is at least as good the square root of the number of video frames, although more work is needed to completely quantify this. Real-time wavelet based video processing and compression technology recently developed under the DoD WaveNet program offers an exciting opportunity to more fully investigate our proposed method.*

1 Introduction

Proximity sensing involves detecting the presence of nearby objects in a system's environment. In radar or sonar sensors, this includes not only the transducing of electromagnetic or acoustic energy to electrical signals, but also the processing of these signals in order to extract useful information. The sensor signals may also be images, for example those from synthetic aperture radar. In such cases we can extract information in both space and time.

In this paper we propose a novel wavelet video based method of processing proximity sensor signals. In proximity sensing, interesting signal structures are usually localized. Wavelet representations are therefore ideal, since they have both spatial and temporal localization. The proximity-sensing problem could then be seen as the recognition of any patterns among the time varying wavelet transform coefficients of the sensor signal that may indicate the presence of objects.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
OCT 1998		**00-00-1998 to 00-00-1998**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Proximity sensing with wavelet generated video	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Surface Warfare Center Dahlgren Division,Systems Research and Technology Department,17320 Dahlgren Road,Dahlgren,VA,22448-5100**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Journal of Electronic Imaging, 7(4), October 1998

14. ABSTRACT

In this paper we introduce wavelet video processing of proximity sensor signals. Proximity sensing is required for a wide range of military and commercial applications, including weapon fuzing, robotics, and automotive collision avoidance. While our proposed method temporarily increases signal dimension, it eventually performs data compression through the extraction of salient signal features. This data compression in turn reduces the necessary complexity of the remaining computational processing. We demonstrate our method of wavelet video processing via the proximity sensing of nearby objects through their Doppler shift. In doing this we perform a continuous wavelet transform on the Doppler signal, after subjecting it to a time varying window. We then extract signal features from the resulting wavelet video, which we use as input to pattern recognition neural networks. The networks are trained to estimate the time varying Doppler shift from the extracted features. We test the estimation performance of the networks, using different degrees of nonlinearity in the frequency shift over time and different levels of noise. We give the analytical result that the signal -to-noise enhancement of our proposed method is at least as good the square root of the number of video frames although more work is needed to completely quantify this. Real-time wavelet based video processing and compression technology recently developed under the DoD WaveNet program offers an exciting opportunity to more fully investigate our proposed method.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**19**	

In our scheme we place a window of fixed width over the incoming signal, so as to localize the processing near the present time. We then perform a continuous wavelet transform on the signal within the window, resulting in an image of the transform. As the signal window then moves forward in time, the corresponding sequence of transform images forms a video. From this wavelet transform video, we then extract features as input to pattern recognition algorithms such as artificial neural networks.

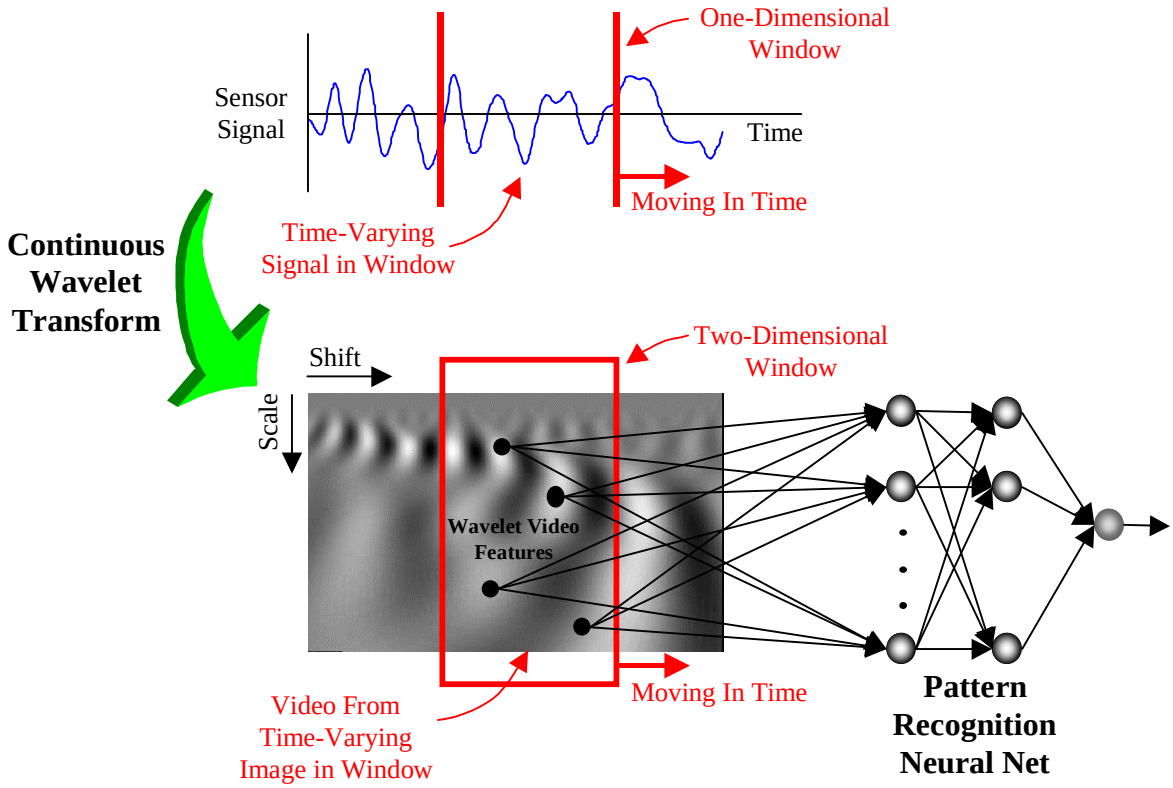


Fig. 1 Generation of wavelet transform video from time varying proximity sensor signal, to provide neural network features. Temporary expansion of dimensionality allows us to extract salient features, leading to reduced computational complexity.

The processing of video data in real time is considered to be somewhat impractical given the current state of technology. As such the utility of such processing in real-world applications would seem to be limited. However, recent developments at Trident Systems, Incorporated have made available real-time wavelet processing of video, in the form of the WaveNet technology[1]. Also, in the future a variety of fast architectures for computing wavelet transforms are likely to be developed.

Besides, in our proposed scheme, wavelet processing is not a particular computational hindrance, but rather allows salient features to be extracted via the wavelet coefficients. Because of the quality of the wavelet features, it is likely that fewer inputs are needed for pattern recognition. In this sense our scheme could be considered to be a form a data compression. In particular, it seems to be a form of data compression that is ideal for pattern recognition.

The multiresolution nature of wavelets also allows us to explore the tolerance of imprecision in the processing of signals. This provides the freedom to tailor the design of the sensor to the resolution requirements of

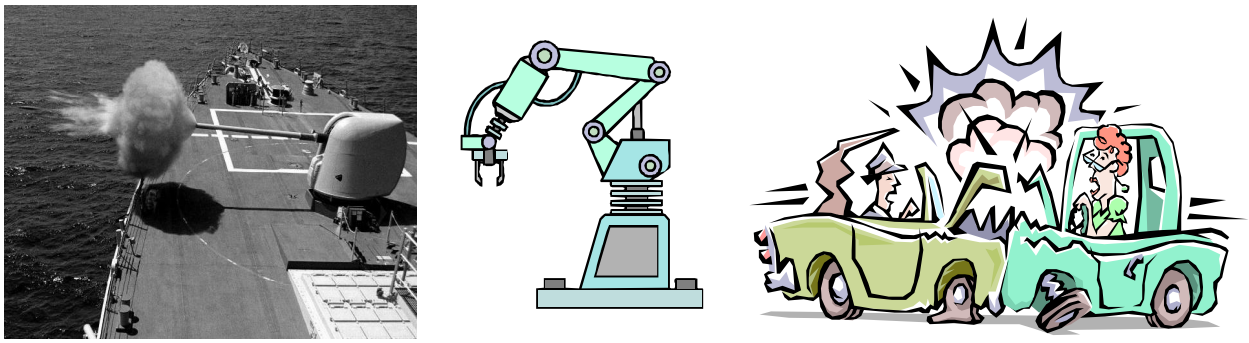
the signals being processed. This tolerance of imprecision is in the spirit of fuzzy logic, but in this case the imprecision is in the scale of the signal structures rather than in the membership of sets.

The important idea is that useful information in signals is generally found at the larger scales (lower frequencies). The less useful, smaller scale signal structures can therefore be disregarded. Neglecting unnecessary details allows a reduction in the amount of data to be processed. This in turn reduces the complexity of the processing, leading to improvements in processing time, system size, and system cost.

This reduction of data through the explicit use of scale is a powerful form of data compression. While there are several other strategies for data compression, this one has the advantage of being based on the extraction of signal features. Through wavelet transform time integration, a single coefficient provides the correlation between the signal and a wavelet at a particular scale and time shift. Wavelets are known to provide good signal features for pattern recognition algorithms such as artificial neural networks. Indeed, natural sensors such as eyes and ears carry out wavelet-type processing.

The continuous wavelet transform effectively increases the dimensionality of the signal representation from one to two. While this might cause some concern at first glance, it is really not a problem. The reason is that the wavelet representation will be used to extract signal features only. Thus the pattern recognition neural networks need not suffer from the "curse of dimensionality." After all, the extracted features are of a single dimension only, so that the increase in dimensionality is only temporary. Indeed, because of the high quality of wavelet features, it is quite possible that fewer features will be needed, and that recognition performance will be improved.

Many systems have the need to sense the proximity of objects in their environment. For example, one of the first applications of radar was as a proximity sensor for military fuzes[2]. Proximity sensing is also widely applied in manufacturing automation and robotics. More recently, there has also been a strong interest in proximity sensing for automobile collision avoidance.



(a) (b) (c)

Fig. 2 Example applications of proximity sensing: (a) detection of targets for military fuzing, (b) manufacturing automation and robotics, and (c) automobile collision avoidance.

In the next section, we introduce an important type of signal for proximity sensing, namely, the Doppler signal. We go on to describe how the Doppler signal can be used to detect the presence of nearby objects. In Section 3, we show the continuous wavelet transform representation of sensor signals, and contrast it with other wavelet and Fourier representations. We also show how fast wavelet denoising algorithms can dramatically improve signal quality, which leads to enhanced recognition performance. Then, in Section 4 we demonstrate how

features extracted from wavelet-generated video can be used along with neural networks to improve proximity sensing. Finally, in Section 5 we summarize our work and draw conclusions.

2 Doppler Signals for Proximity Sensing

Active proximity sensors such as radar, acoustic, and ladar detect objects by first emitting energy waves, then receiving the signal from the reflected waves. These sensors can then process the echo signal in order to gain information about the objects. An important type of information for such sensors is the change in frequency of the echo signal relative to the emitted signal. The frequency shift is proportional to the relative velocity between the object and the sensor. This is the well-known Doppler effect. A familiar example of the Doppler effect is the noticeable change in whistle pitch as a train passes by.

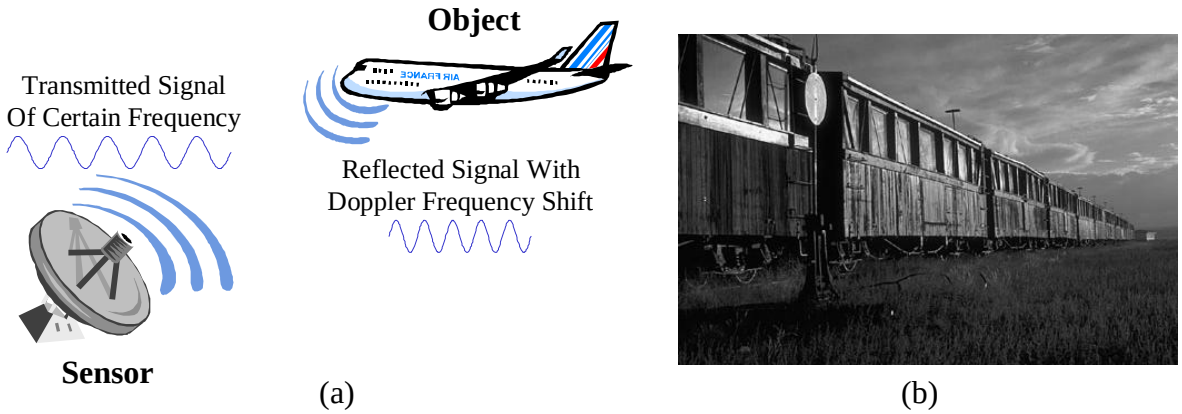


Fig. 3 Doppler effect: (a) Doppler induced change in frequency between transmitted and reflected signals for active sensors, and (b) familiar example of Doppler effect is change in pitch as train passes by.

2.1 *The Doppler Shift and Its Relation to Relative Velocity*

It is well known in electromagnetics, acoustics, and optics that if either the source or observer of an oscillating wave is in motion, the oscillation frequency appears to shift. This shift is the Doppler effect. In the case of proximity sensing, the source and observer are both located in the sensor. The Doppler effect then arises from the relative motion between the sensor and the sensed object.

Let us derive the Doppler frequency shift and the corresponding relative velocity between sensor and object. Assume that the distance between the sensor and object is R. The total number of radiation wavelengths λ over the transmitted and reflected paths is then $2R/\lambda$. Since one wavelength λ corresponds to an angular phase of 2π, the total phase ϕ is $4\pi R/\lambda$. The rate of change in ϕ with time t is the angular Doppler frequency ω_d, which is then

$$\omega_d = 2\pi f_d = \frac{d\phi}{dt} = \frac{4\pi}{\lambda}\frac{dR}{dt} = \frac{4\pi v_r}{\lambda}. \qquad (1)$$

Here f_d is the Doppler frequency shift and v_r is the relative velocity of the object with respect to the sensor. The Doppler frequency shift f_d then becomes

$$f_d = \frac{2v_r}{\lambda} = \frac{2v_r f_0}{c}, \qquad (2)$$

where f_0 is the transmitted frequency and c is the velocity of radiation propagation. The relative velocity v_r corresponding to this frequency shift is then

$$v_r = \left(\frac{c}{2}\right)\frac{f_d}{f_0}. \qquad (3)$$

Thus we can calculate the relative velocity v_r between the sensor and an object from measurements of the Doppler frequency shift f_d.

2.2 *Direction to Sensed Object from Doppler Shift*

We can also calculate the direction to the object from measurements of the Doppler frequency shift. Figure 4 shows a sensor and an object to be sensed. We assume they are approaching at a constant velocity $\vec{v}$, with the origin fixed at the sensor. The relative velocity v_r can be written as

$$v_r = |\vec{v}|\cos\theta = v\cos\theta, \qquad (4)$$

where θ is the angle toward the object. The angle θ is then calculated as

$$\theta = \cos^{-1}\left(\frac{v_r}{v}\right). \qquad (5)$$

This calculation requires a value of the magnitude $v = |\vec{v}|$ of the approach velocity $\vec{v}$, which can be estimated by measuring the relative velocity v_r at a distance sufficiently larger than the closest approach distance α. In fact, if we have reliable measurements of v and the time of closest approach between sensor and object, we can calculate α as

$$\alpha = \frac{-vt}{\cot(\theta)}, \qquad (6)$$

with $t = 0$ at the time of closest approach.

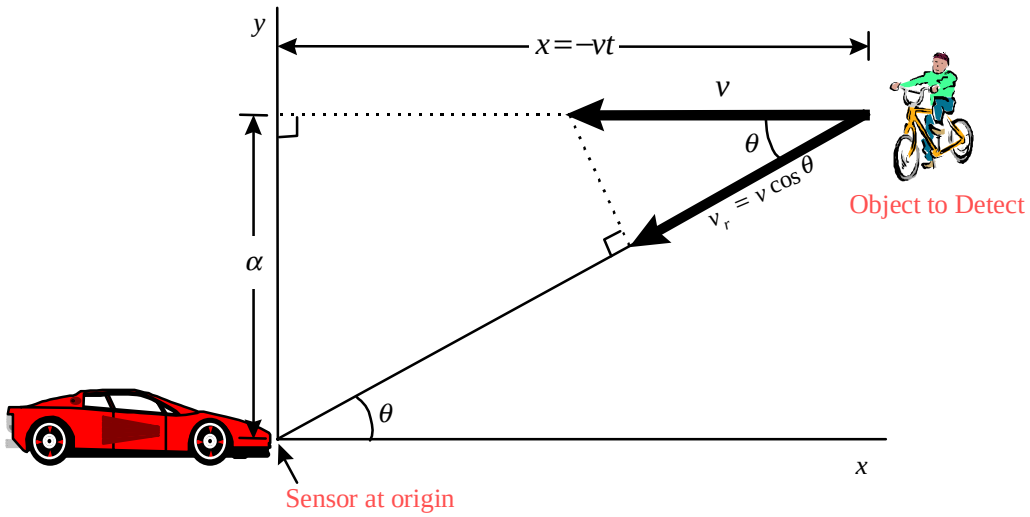


Fig. 4 Geometry for sensor and object over time.

2.3 *Utility of Doppler Shift in Proximity Sensing*

While the kinematics we just described are somewhat idealized, the general behavior does hold true in most cases. At sufficiently large distances between the sensor and an object, both the relative velocity and the corresponding Doppler frequency shift asymptotically approach a constant. As the object later passes near the sensor, the relative velocity decreases, with the frequency shift decreasing proportionately. The nearer the object passes the sensor, the more nonlinear is the change in Doppler frequency shift over time. This general behavior is shown in Figure 5.

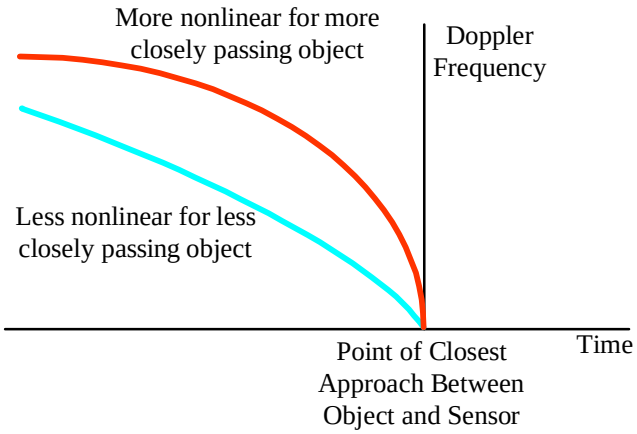


Fig. 5 Change in Doppler frequency over time as an object passes near proximity sensor.

Proximity sensors can take advantage of this behavior in order to gain information about sensed objects. For example, the approach of Doppler frequency shift to zero indicates that the object is at its closest distance from the sensor. Another example can be taken from the military weapon fuzing problem. Here the optimal value of θ for fuzing is known to be

$$\theta = \tan^{-1}\left(v_{\text{frag}} / v\right), \qquad (7)$$

where v_{frag} is the velocity of the warhead fragments. Note that this is independent of the closest app roach distance α. The fuzing problem is then to estimate the Doppler shift over time, and to detonate when the shift reaches its optimal value.

3 Wavelet Representation and Denoising of Doppler Signals

The Fourier transform is the cornerstone of signal processing. However, since it lacks time localization, it is less suited to the processing of signals whose frequencies change over time. The time-dependent (or windowed) Fourier transform localizes time by doing the transform over a window, which shifts in time. Unfortunately, the width of the window is fixed over the entire transform, which causes problems in the high-frequency limit[3].

In contrast, a wavelet transform has a window whose bandwidth varies in proportion to the center frequency of the wavelet. This is the so-called constant-*Q* property from electrical engineering. The result is that the wavelet transform performs time-scale processing rather than time-frequency processing. Also, wavelet transforms allow more freedom in the choice of basis, so that the basis functions can be better matched to the shape of the signal.

The wavelet transform provides the local scale of the signal over time, which for Doppler signals is the local period or inverse of frequency. Wavelet representations of the Doppler signal are particularly necessary in the case of closely passing objects, for which the change in frequency is more abrupt. These representations are also convenient when the signal is embedded in nonstationary noise.

In the remainder of this section, we first explain the advantages of the continuous wavelet transform over the discrete transform for pattern recognition problems. We then formally introduce the continuous wavelet transform, and show how it generates a two-dimensional representation (image) for proximity sensing Doppler signals. We perform the transform with the real-valued Morlet wavelet[4], which is well matched to the Doppler signals of interest. We also contrast this transform to the time-dependent Fourier transform with a Gabor window[5]. To improve performance for noisy signals, we apply a fast wavelet-based denoising algorithm.

3.1 *Advantages of Continuous Wavelet Transform for Pattern Recognition*

Mallat's multiresolution analysis[6] leads to discrete orthogonal wavelets at dyadic scales and shifts, implemented via the efficient pyramid algorithm. These discrete wavelets have been successful in many applications, particularly data compression. However, discrete wavelets have limited utility for pattern recognition problems. This is because interesting signal structures are not constrained to follow such power-of-two patterns. In particular, discrete wavelet transform coefficients are shift-variant, which in general causes problems for pattern recognition.

In contrast, the continuous wavelet transform has coefficients at all scales and shifts, not just dyadic ones. The continuous transform therefore has the desirable property of shift invariance. Another advantage of continuous wavelets is that they have less stringent requirements for admissibility, which allows a wider choice of basis functions. They also have the possibility of being basis functions for adaptive wavelet networks.

Through the inclusion of all scales and shifts, the continuous wavelet transform effectively increases the dimensionality of the signal representation. That is, the representation is made to be a function of two variables rather than one. We note that the discrete wavelet transform introduces no such increase in dimensionality, since the number of transform coefficients is the same as the number of signal sample points. This is because the discrete wavelet transform employs an orthonormal basis rather than an overcomplete frame.

However, the fact that we are using the continuous wavelet transform coefficients merely for feature extraction means that we need not be plagued by the curse of increased dimensionality. In particular, the goal is to use only the relatively few coefficients that provide the best features. In fact, the use of such high quality features may well mean that fewer inputs are ultimately needed for pattern recognition. Of course, these high quality features are also likely to improve the performance of the neural networks. In this sense, the temporary increase in dimensionality could actually improve compression quality, at least when measured with respect to recognition performance.

If we disregard the issue of dimensionality, it might still be argued that computation of the discrete wavelet transform is faster, which has complexity $O(n)$. However, a continuous wavelet transform implemented via the fast Fourier transform has complexity $O(n \log n)$, which is still quite acceptable for many applications. Also, a continuous wavelet transform has the potential for massive parallelism.

3.2 *Continuous Wavelet Transform and Contrast to Gabor Transform*

The continuous wavelet transform[4] $F_w(a,b)$ of a signal $f(t)$ is given by

$$F_w(a,b) = a^{-1/2} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a} \right) dt, a > 0 . \qquad (8)$$

Here a and b are scale and shift parameters, respectively. A necessary and sufficient condition for Eq. (8) to be invertible is that $\psi(t)$ satisfies the wavelet admissibility condition

$$\int_{-\infty}^{\infty} |\Psi(\omega)|^2 |\omega|^{-1} d\omega < \infty , \qquad (9)$$

where $\Psi(\omega)$ is the Fourier transform of $\psi(t)$. If $\psi(t)$ has reasonable smoothness and decay at infinity, which is usually the case, the admissibility condition can be written as

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 . \qquad (10)$$

Under certain conditions, it is possible to reconstruct $f(t)$ from samples of $F_w(a,b)$ taken on a hyperbolic lattice. The collection of wavelet functions $\psi\left(\frac{t-b}{a}\right)$ over this lattice is then said to constitute a frame. A frame, in contrast to a basis, is an overcomplete set. This redundant representation allows more flexibility in the choice of inputs to pattern recognition neural networks. In particular, we are not constrained to the power-of-two scales characteristic of the discrete wavelet transform.

We choose for $\psi(t)$ the real part of the Morlet wavelet[7], which is

$$\psi(t) = \mathrm{Re}\left(e^{-i\omega_0 t} e^{-t^2/2}\right) = \cos(\omega_0 t) e^{-t^2/2}, \qquad (11)$$

with $\omega_0 = \pi\sqrt{2/\ln 2} = 5.336$, which is a standard value. The real Morlet wavelet is a Gaussian-modulated sinusoid, which is well suited to processing sinusoidal Doppler signals. The wavelet transform with the real Morlet is similar to the time-dependent cosine Fourier transform with a Gabor[3] (Gaussian-shaped) window. This type of Fourier transform is also called the Gabor transform, and is given by

$$F(\omega, b) = \int_{-\infty}^{\infty} f(t)\cos(\omega t) e^{-(t-b)^2/2} dt\,. \qquad (12)$$

For comparison, we can write the wavelet transform in Eq. (8) as

$$F_w(\omega', b) = \int_{-\infty}^{\infty} f(t)\cos[\omega'(t-b)] e^{-\left(\frac{t-b}{\omega_0/\omega'}\right)^2/2} dt\,, \qquad (13)$$

where $\omega' = \omega_0 / a$. For the Gabor transform in Eq. (12), the width of the window W_G, given by

$$W_G = \exp\left[-\frac{1}{2}(t-b)^2\right] \qquad (14)$$

remains fixed. However, for the wavelet transform in Eq. (13) the window width W_W, given by

$$W_W = \exp\left[-\frac{1}{2}\left(\frac{t-b}{\omega_0/\omega'}\right)^2\right] \qquad (15)$$

varies inversely with the frequency $\omega' = \omega_0 / a$. Thus the frequency bandwidth of the wavelet window varies in proportion to ω', through the inverse scaling property of Fourier conjugate variables. Also, the cosine term $\cos[\omega'(t-b)]$ for the wavelet transform shifts in time along with the window, through the shift parameter b. In contrast, for the Gabor transform only the window shifts in time, and the cosine term remains fixed.

3.3 *Fast Wavelet Denoising*

To improve performance for noisy Doppler signals, we apply Donoho's $O(n)$ wavelet denoising algorithm[8]. The algorithm first does the discrete wavelet transform with Mallat's pyramid algorithm[6]. The pyramid algorithm computes the transform for some J dyadic levels of scale, resulting in vectors of detail and smooth wavelet coefficients $\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_{J-1}, \mathbf{d}_J, \mathbf{s}_J$. The algorithm then shrinks the detail coefficients for scales $j \le J-1$ to obtain $\tilde{\mathbf{d}}_1, \tilde{\mathbf{d}}_2, \ldots, \tilde{\mathbf{d}}_{J-1}$. Here the $\tilde{\mathbf{d}}_j$ are

$$\tilde{\mathbf{d}}_j = \delta_{\lambda_j \sigma_j}(\mathbf{d}_j)\,, \qquad (16)$$

where $\delta_{\lambda\sigma}(x)$ is a nonlinear threshold shrinkage function given by

$$\delta_{\lambda\sigma}(x) = \begin{cases} 0 & \text{if } |x| \le \lambda\sigma \\ \text{sign}(x)(|x| - \lambda\sigma) & \text{if } |x| > \lambda\sigma. \end{cases} \quad (17)$$

This threshold shrinkage function is shown in Figure 6.

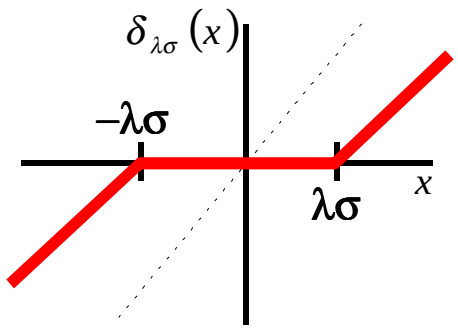


Fig. 6 Nonlinear threshold shrinkage function for wavelet denoising.

The threshold shrinkage function $\delta_{\lambda\sigma}(x)$ is parameterized by a threshold λ and an estimate of the standard deviation of the noise σ. We use a universal threshold[8] $\lambda_j = \sqrt{2\log(N)}$, where N is the number of data samples. For σ we use the median absolute deviation, which is a robust estimation of standard deviation. Finally, the denoising algorithm computes the inverse discrete wavelet transform using the new coefficients $\tilde{\mathbf{d}}_1, \ldots, \tilde{\mathbf{d}}_{J-1}, \mathbf{d}_J, \mathbf{s}_J$. This results in a non-parametric estimate of the signal without the noise. The entire wavelet denoising algorithm is shown in Figure 7.

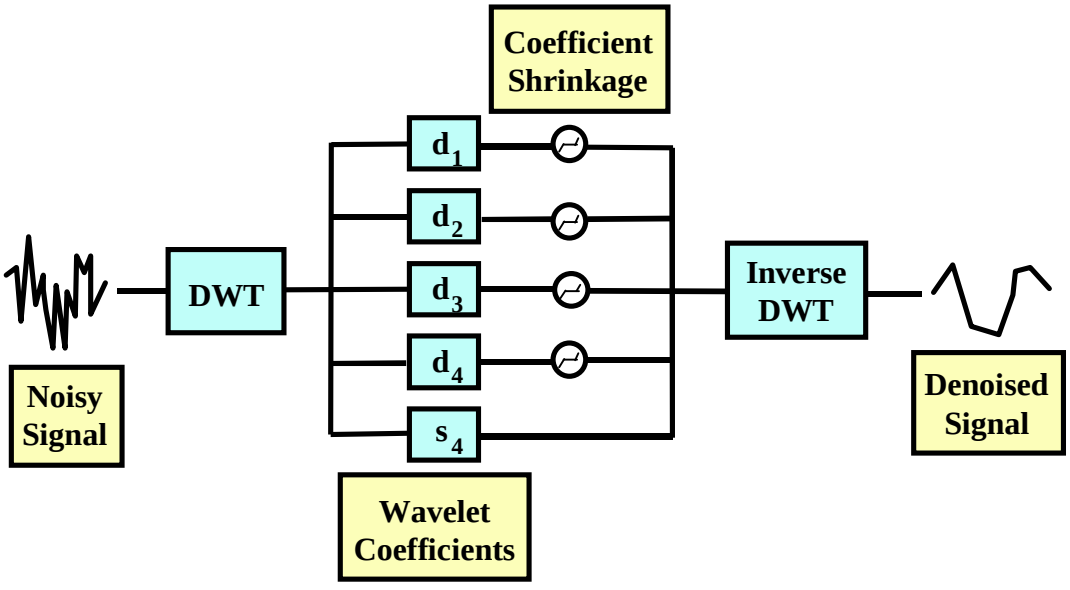


Fig. 7 Wavelet denoising algorithm.

For the discrete wavelet transform in the denoising algorithm, we apply a super-Haar wavelet[9], which is a linear superposition of shifted Haar wavelets. The super-Haar scaling function $\phi(t)$ is given by

$$\phi(t) = \sum_k s_k \phi_H(t-k), \qquad (18)$$

where s_k are integer coefficients and $\phi_H(t)$ is the Haar scaling function[10], given by

$$\phi_H(t) = \begin{cases} 1, & t \in [0,1) \\ 0, & t \notin [0,1). \end{cases} \qquad (19)$$

We apply the particular super-Haar in which $s_k = [1,2,2,1]$.

3.4 *Simulations*

Figure 8 shows pure, noisy, and denoised versions of a simulated Doppler signal. The closest-approach distance α is such that the change in frequency is nearly linear over time. We assume that the sinusoid amplitude is constant over time, which is appropriate over the short distances applicable to proximity sensing.

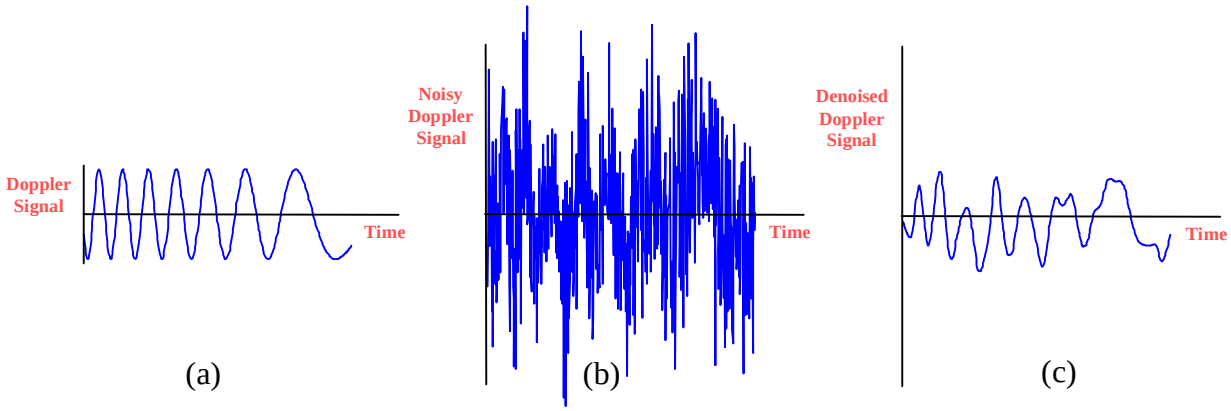


Fig. 8 Doppler signals: (a) pure, (b) noisy, and (c) denoised.

Figure 9 shows the continuous wavelet transforms of the three signals in Figure 8. Figure 10 shows the same three transforms, using a surface plot rather than a grayscale image. The wavelet transforms show the increase in local signal scale over time. In this case the increasing signal scale is the increasing period of the frequency-modulated sinusoid.

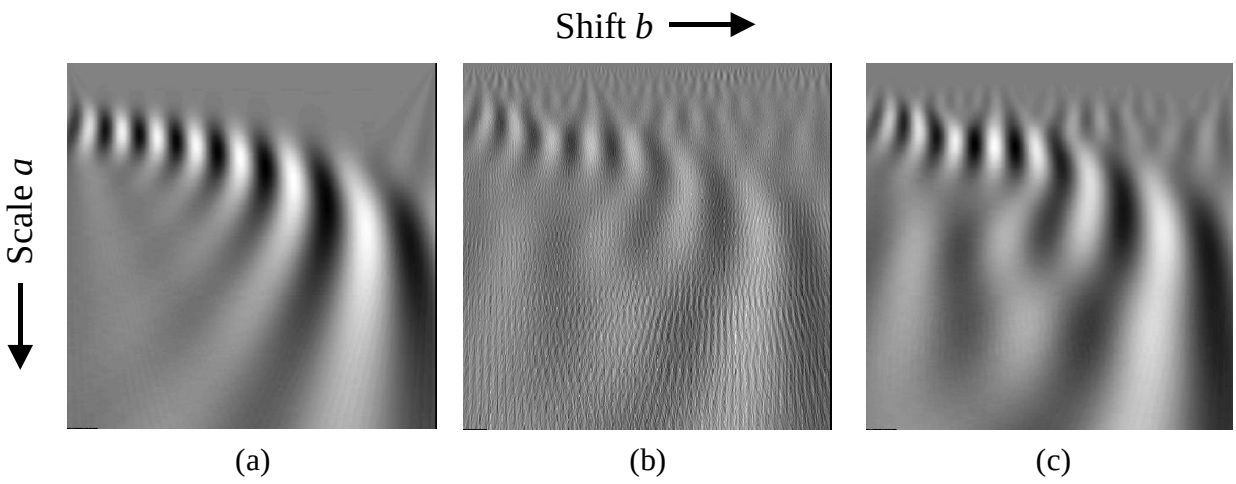


Fig. 9 Continuous wavelet transforms (grayscale image): (a) pure, (b) noisy, and (c) denoised signals.

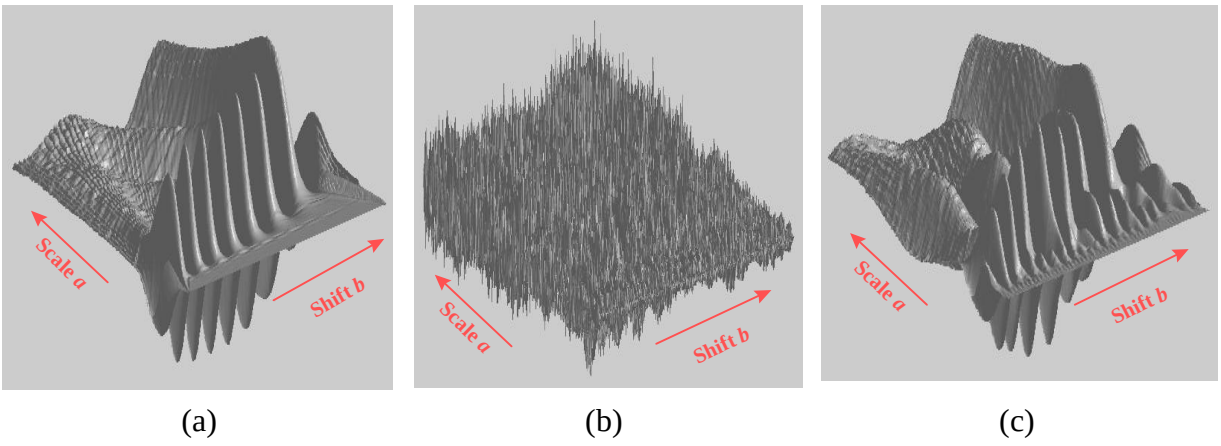


Fig. 10 Continuous wavelet transforms (surface plots): (a) pure, (b) noisy, and (c) denoised signals.

The time-scale structure of the Doppler signal is visually apparent to some extent in the transform of the noisy signal. However, if samples of the noisy transform were used as neural network inputs for proximity sensing, the high-frequency fluctuations would result in poor performance. These fluctuations are largely removed by the wavelet denoising, which will result in much improved performance.

4. Proximity Detection with Wavelet Video Features and Neural Networks

The continuous wavelet transform correlates a Doppler signal with time-localized wavelets at various scales and shifts. It gives the change in local signal scale over time, which in this case is the Doppler period or inverse frequency. When a moving window is placed on the incoming Doppler signal and the windowed signal is wavelet transformed, the corresponding time-varying transform imagery constitutes video. Samples of this wavelet-generated video over time then form signal features for pattern recognition neural networks. These networks are then trained to extract the Doppler frequency shift over time. This frequency shift is critical information for proximity sensing.

The continuous wavelet transform constitutes a frame rather than a basis. Such a redundant representation allows more flexibility in the selection of signal features. In terms of the most efficient signal representation, these features should be orthogonal. However, such a representation in which the features are completely independent is less robust with respect to noise immunity and fault tolerance. The search for the best representation is therefore a tradeoff between redundancy and robustness[11].

We extract the Doppler shift with feedforward multilayer neural networks, known as multilayer perceptrons[12]. After computing the continuous wavelet transform of the denoised Doppler signal, we sample the transform coefficients to provide inputs for the multilayer perceptrons. The networks are trained with the Levenberg-Marquardt rule[13] to provide the Doppler shift at a given time. This rule is a powerful generalization of gradient descent that employs an approximation of Newton's method. It is much faster than standard gradient descent algorithms such as backpropagation, although it does require more memory.

In the remainder of this section, we first describe the architecture and training algorithm we employ for the pattern recognition neural networks. We then show simulations that demonstrate the improvement offered by signal features taken from wavelet-generated video. Finally, using wavelet-generated video features, we show pattern recognition performance for the estimation of time-varying Doppler shift. We show this performance for different degrees of nonlinearity in the shift over time, as well as performance for different levels of noise.

4.1 *Architecture and Training for Pattern Recognition Neural Networks*

Figure 11 shows the neural network architecture we employ for Doppler frequency estimation. The network is comprised of 3 layers of artificial neurons: an input layer, a middle or hidden layer, and an output layer. Signals flow forward through the network, that is, from input layer to hidden layer to output layer. This architecture is known as a multilayer feedforward network, or multilayer perceptron.

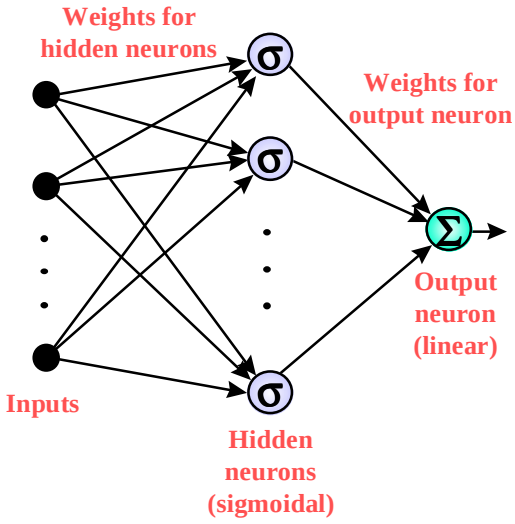


Fig. 11 Neural network architecture for proximity sensing pattern recognition.

The input neuron layer in Figure 11 performs no processing; it merely provides means for coupling the input vectors to the hidden layer. The neurons in the middle layer sum the weighted network inputs, along with an internal bias for each neuron, then apply the nonlinear sigmoidal activation function

$$\sigma(v_j) = \tanh\left(\frac{v_j}{2}\right) = \frac{\left(1 - e^{-v_j}\right)}{\left(1 + e^{-v_j}\right)}, \qquad (20)$$

where v_j is the weighted sum for neuron j. This sigmoidal nonlinearity limits the neuron outputs to (-1,1). The single output neuron computes the weighted sum of the outputs of the hidden neurons, along with its internal bias, without applying the sigmoidal function.

The architecture in Figure 11 is known to be a universal function approximator[12], that is it can represent an arbitrary function arbitrarily well, given a sufficient number of neurons in the hidden layer. The particular function mapping that the network performs is determined by the values of the weights between neuron layers and the internal neuron biases.

Various learning algorithms exist for computing the network weights and biases for a given problem. The most popular learning algorithm is backward error propagation[12], which attempts to minimize the squared error of the network through gradient descent in weight space. We can define the error signal for neuron j as

$$e_j(n) = d_j(n) - y_j(n), \qquad (21)$$

where n indexes the training vectors, $d_j(n)$ is the desired response for neuron j, and $y_j(n)$ is the actual output of neuron j. The instantaneous value of the sum of squared errors $\frac{1}{2}e_j^2(n)$ over all neurons in the output layer of the network can then be written as

$$\mathsf{E}(n) = \frac{1}{2}\sum_{j \in C} e_j^2(n), \qquad (22)$$

where the set C includes all neurons in the output layer and N is the number of vectors in the training set. The squared error averaged over all training vectors is then

$$\mathsf{E}_{\text{av}} = \frac{1}{N}\sum_{n=1}^{N} \mathsf{E}(n). \qquad (23)$$

The average squared error E_{av} constitutes a cost function that is to be minimized. It is minimized approximately by iteratively reducing $\mathsf{E}(n)$ for each training vector. The correction $\Delta w_{ji}(n)$ to be applied to weight $w_{ji}(n)$ is then defined by the delta rule

$$\Delta w_{ji}(n) = -\eta \frac{\partial \mathsf{E}(n)}{\partial w_{ji}(n)}, \qquad (24)$$

where η is a parameter that determines the rate of learning. The minus sign in Eq. (24) results in gradient descent in weight space, that is weights are moved in the opposite direction of the error gradient.

We apply a powerful generalization of backward error propagation known as the Levenberg-Marquardt weight update rule[13]. This rule can be written in matrix notation as

$$\Delta \mathbf{W} = \left(\mathbf{J}^{\mathrm{T}} \mathbf{J} + \mu \mathbf{I} \right)^{-1} \mathbf{J}^{\mathrm{T}} \mathbf{e} , \qquad (25)$$

where $\Delta \mathbf{W}$ is the matrix of weight updates, $\mathbf{e}$ is the error vector, and $\mathbf{J}$ is the Jacobian matrix of derivatives of each error to each weight. If the parameter μ is very large, Eq. (23) approximates gradient descent, while if μ is small it becomes the Gauss-Newton method.

The Gauss-Newton method is faster and more accurate near an error minimum. The idea is therefore to shift towards Gauss-Newton as quickly as possible. The parameter μ is thus decreased after each successful step, and increased only when a step increases the error. The Levenberg-Marquardt update rule is known to train networks much more quickly than standard backward error propagation. However, it does require more memory, usually a factor of $C * N$ more, where C is the number of output neurons and N is the number of training vectors.

4.2 *Signal Features from Wavelet-Generated Video*

We now demonstrate the improvement in pattern recognition performance that wavelet-generated video features can provide. We begin by showing how wavelet transform features outperform both time-domain and frequency-domain ones for classifying signals according to frequency. In particular, we test the ability of pattern recognition neural networks to classify signals as either being either above or below a certain threshold frequency, in the presence of noise.

Figure 12 shows pattern recognition performance using 3 different signal representations for neural network input: wavelet transform coefficients, time-domain samples, and Fourier transform coefficients. A variety of frequencies were used for the test signals, equally distributed about the threshold frequency. The networks were trained to determine whether the signal frequencies were below (output of zero) or above (output of one) the threshold. Because of the binary nature of this experiment, the networks were made to have sigmoidal rather than linear activation functions. The noise was white Gaussian, with signal-to-noise ratio of –1 dB.

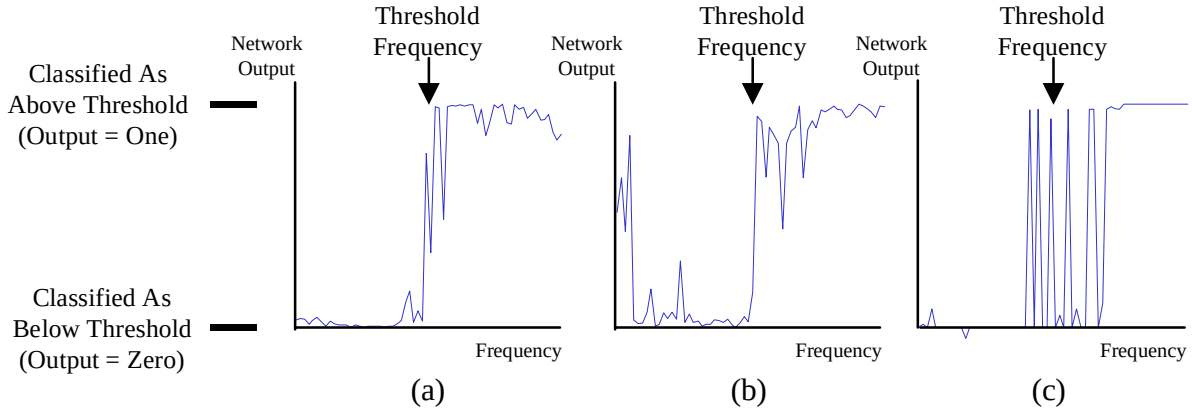


Fig. 12 Frequency classification performance using (a) wavelet, (b) time-domain, and (c) frequency-domain features.

We see from Figure 12 that classification performance for wavelet features is better than for either time-domain or frequency-domain features. In particular, for the time-domain features there are many misclassifications at the lowest frequencies, and at frequencies just above the threshold. Also, for the frequency-domain features there are many misclassifications near the threshold.

If we look more carefully at Figure 12, we see that at the highest frequencies, and at frequencies just below the threshold, performance is slightly better for time-domain features than for wavelet features. Also, at the lowest and highest frequencies, performance is slightly better for frequency-domain features than for wavelet features. Interestingly, it appears that the wavelet transform has formed a compromise between the time and frequency domains in which overall classification performance is improved.

Now that we have demonstrated the superior frequency classification performance of wavelet features, we can investigate which wavelet transform coefficients might provide the best features for estimating time varying Doppler shift. One fundamental issue is whether to sample from a single time shift of the transform, or to sample over multiple shifts. While sampling from a single shift completely localizes time, which is advantageous in some applications, sampling over multiple shifts gives additional information that may improve estimation performance. Also, sampling over multiple shifts provides a degree of redundancy that will likely improve performance for noisy signals.

As a test of single-shift versus multiple-shift features, we used each type of feature as input to a pattern recognition neural network. For single-shift features, we used 32 samples of the continuous wavelet transform of the Doppler signal, taken over various scales at a single time shift. For multiple-shift features, we used 16 samples of the transform at the original time shift, and 16 more samples at an additional time shift of 6. Figure 13 shows the sampling scheme for the multiple-shift case.

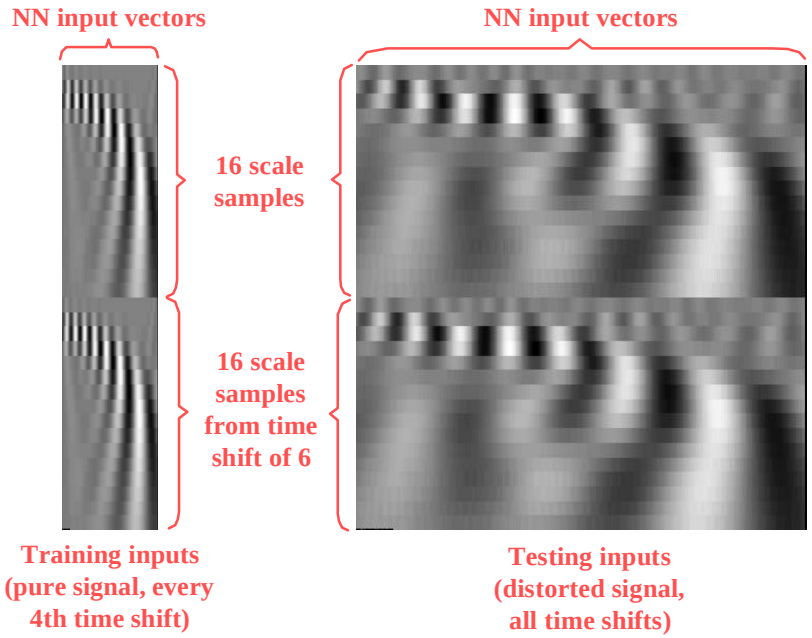


Fig. 13 Samples of wavelet-generated video for inputs to pattern recognition neural networks.

We trained the neural networks with samples of transforms of pure Doppler signals, sampling only every 4^{th} time shift of the transform. For training outputs, we supplied the known instantaneous frequency of the pure

signals for each time shift. Thus the networks were trained to estimate the instantaneous frequency of the Doppler signals, given samples of their wavelet transform.

After training for frequency estimation, we tested the networks with transforms of denoised versions of noisy Doppler signals. The networks were tested for every time shift of the transform, with a noise level of –2 dB. Figure 14 shows the test results. It is obvious that performance is much better for the case of sampling wavelet coefficients over multiple time shifts.

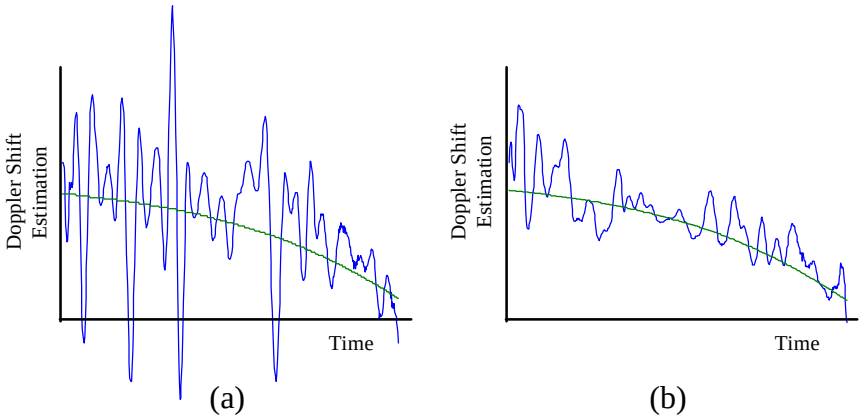


Fig. 14 Estimation performance for time varying Doppler shift: (a) sampling at single scale of wavelet transform, and (b) sampling from wavelet generated video. Smooth lines show true Doppler shift over time.

We point out that sampling over multiple time shifts versus sampling at only a single shift constitutes true image sampling, since both scale and shift variables (two dimensions) are sampled. When the Doppler signal window is then moved forward in time, we have sequence of images over time, that is, we have video. In our simulations, we sample from this wavelet-generated video. In particular, for each wavelet transform image in the sequence that forms video, the samples provide an estimate, via neural networks, of the instantaneous Doppler shift corresponding to the image.

4.3 *Performance for Time Varying Doppler Shift Estimation*

We have just demonstrated the effectiveness of features extracted from wavelet-generated video. We now test the pattern recognition performance of such features in the estimation of time varying Doppler shift from noisy sensor signals. In particular, we sample the wavelet video as shown in Figure 13, use the samples as neural network inputs, and then train the networks to estimate the Doppler shift. We test performance for different degrees of nonlinearity in the Doppler shift over time, as well as for different levels of noise.

Figure 15 shows network test results for various signal-to-noise ratios, where the Doppler signal frequency decreases nearly linearly over time, corresponding to a relatively large closest-approach distance α. The networks were tested for every time shift of the wavelet transform. Since the networks were trained with only every 4^{th} sample, this shows their ability to generalize to other frequencies. Network performance is relatively good, but degrades with decreasing signal-to-noise ratio as would be expected. Figure 16 shows similar network performance for smaller α, which corresponds to a more pronounced nonlinearity in the frequency shift over time.

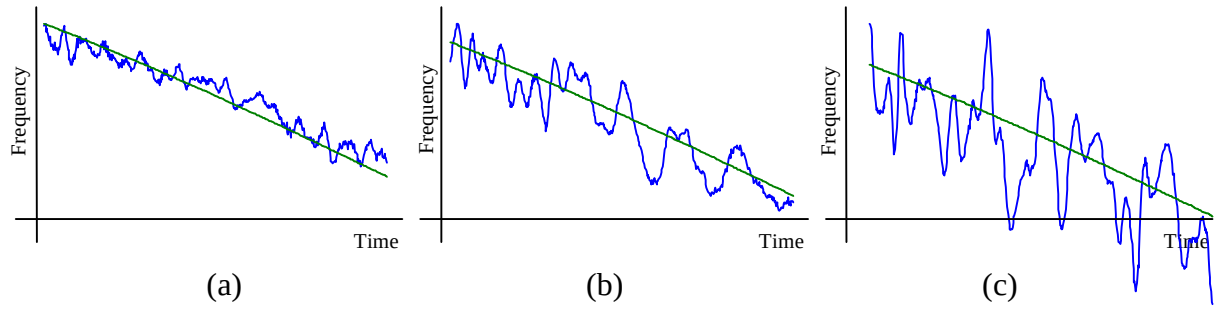


Fig. 15 Estimation of Doppler shift using features from wavelet-generated video and pattern recognition neural networks (nearly linear change in shift over time): (a) signal-to-noise ratio of –0.5 dB, (a) signal-to-noise ratio of –2 dB, (a) signal-to-noise ratio of –4 dB.

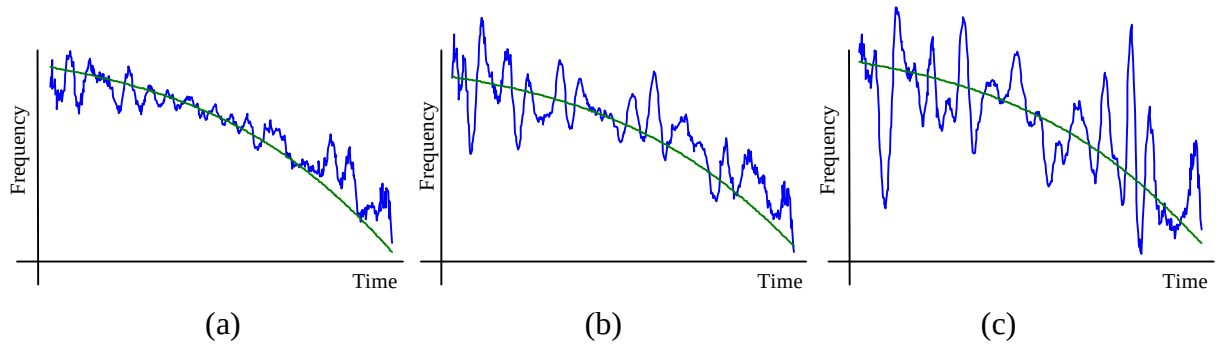


Fig. 16 Estimation of Doppler shift using features from wavelet-generated video and pattern recognition neural networks (nonlinear change in shift over time): (a) signal-to-noise ratio of –0.5 dB, (a) signal-to-noise ratio of –2 dB, (a) signal-to-noise ratio of –4 dB.

Analytically, for a signal in which the frequency content is constant over time, and assuming white noise, the signal-to-noise enhancement of our proposed method of processing is a factor of $\sqrt{N}$, where N is the number of video frames. However, in our experience, for signals with time varying frequency components, or when the noise is nonstationary, an improvement exceeding this $\sqrt{N}$ can be expected. A more detailed analysis is necessary to further quantify this.

5 SUMMARY AND CONCLUSIONS

The value of this paper is to introduce the processing of proximity sensor signals through wavelet generated video. While temporarily increasing signal dimension through a representation in both scale and shift, the method ultimately performs data compression through the extraction of signal features. This reduction of data in turn reduces the overall computational complexity. Moreover, existing hardware and software developed under the DoD WaveNet program can potentially provide a testbed in which to further evaluate this method. Because of the many important military and commercial applications of proximity sensing, it is worthwhile to pursue this work.

We demonstrated our method of video processing by detecting the proximity of objects through their Doppler shift. We placed a time varying window over the Doppler signal, then performed a continuous wavelet transform on the windowed signal. We then extracted signal features from the resulting wavelet video, which we

used as input to pattern recognition neural networks. The networks were then trained to estimate the time varying Doppler shift from the extracted features.

We tested the estimation performance of the networks, for different degrees of nonlinearity in the frequency change over time, and for different levels of noise. We gave analytical results indicating that the signal-to-noise enhancement of our proposed method is better than the square root of the number of video frames, though more work is needed to completely quantify this. Our main purpose at this point is to demonstrate the utility of using wavelets to reduce the computational complexity of video processing, as applied to proximity sensing.

REFERENCES

1. This technology for real-time wavelet based video processing and compression is being developed under the DoD WaveNet program, led by Szu et al.
2. M. Skolnik, *Introduction to Radar Systems*, 2nd edition, McGraw-Hill, 1980.
3. J. Morlet, G. Arens, I. Fourgeau, D. Girad, "Wave Propagation and Sampling Theory," *Geophysics*, Vol. 47, pp. 203-236, 1982.
4. G. Strang, T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge, 1996.
5. A. Oppenheim, R. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, 1989.
6. S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. on Patt. Anal. and Mach. Intel.*, Vol. 11, No. 7, pp. 674-693, 1989.
7. Y. Meyer, Wavelets: Algorithms and Applications, SIAM, 1993.
8. D. Donoho, I. Johnstone, "Adapting to Unknown Smoothness via Wavelet Shrinkage," Technical report, Department of Statistics, Stanford University, 1992.
9. H. Szu, J. Garcia, B. Telfer, M. Raghuveer, "Super-Haar Designs of Wavelet Transforms," *Proc. SPIE Wavelet Applications III*, Vol. 2762, pp. 152-164, 1996.
10. A. Haar, "Zur Theory der Orthogonalen Funktionen-Systeme," *Math. Ann.*, Vol. 69, pp. 331-371, 1910.
11. H. Szu, P. Watanapongse, "Invertible Global Fourier and Local Wavelet Transforms versus Non-Invertible Global Principle Component Analysis and Local Independent Component Analysis," preprint.
12. S. Haykin, *Neural Networks – A Comprehensive Foundation*, Macmillan, 1994.
13. H. Demuth, M. Beale, *MATLAB Neural Network Toolbox User's Guide*, The MathWorks, 1996.